

# Mexican Sign Language Recognition Using Jacobi-Fourier Moments

Francisco Solís<sup>1</sup>, Carina Toxqui<sup>2</sup>, David Martínez<sup>1</sup>

<sup>1</sup>University Center UAEM Teotihuacan Valley, Autonomous University of Mexico State, México

<sup>2</sup>Polytechnic University of Tulancingo, Tulancingo, México

Email: [jfsolisv@uamex.mx](mailto:jfsolisv@uamex.mx), [ctoxqui@upt.edu.mx](mailto:ctoxqui@upt.edu.mx)

Received 13 July 2015; accepted 27 October 2015; published 30 October 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The present work introduces a system for recognizing static signs in Mexican Sign Language (MSL) using Jacobi-Fourier Moments (JFM) and Artificial Neural Networks (ANN). The original color images of static signs are cropped, segmented and converted to grayscale. Then to reduce computational costs 64 JFM were calculated to represent each image. The JFM are sorted to select a subset that improves recognition according to a metric proposed by us based on a ratio between dispersion measures. Using WEKA software to test a Multilayer-Perceptron with this subset of JFM reached 95% of recognition rate.

## Keywords

Mexican Sign Language, Jacobi-Fourier Moments, Digital Image Processing

---

## 1. Introduction

Sign Language Recognition (SLR) is a research field that has grown in recent years; researchers around the world are increasingly interested more in this area. Sign Language (SL) is the main and the most natural form of communication for unhearing community; however, most people interact verbally, leading to the SL mainly limited to the deaf people and closest hearing people that interact with them [1].

SL is not universal; each country or region has its own SL, Mexican Sign Language (MSL), American Sign Language (ASL), Chinese Sign Language (ChSL), Japanese Sign Language (JSL), Persian Sign Language (PSL); to name a few, there are differences between them depending on their uses and customs where they are used [2].

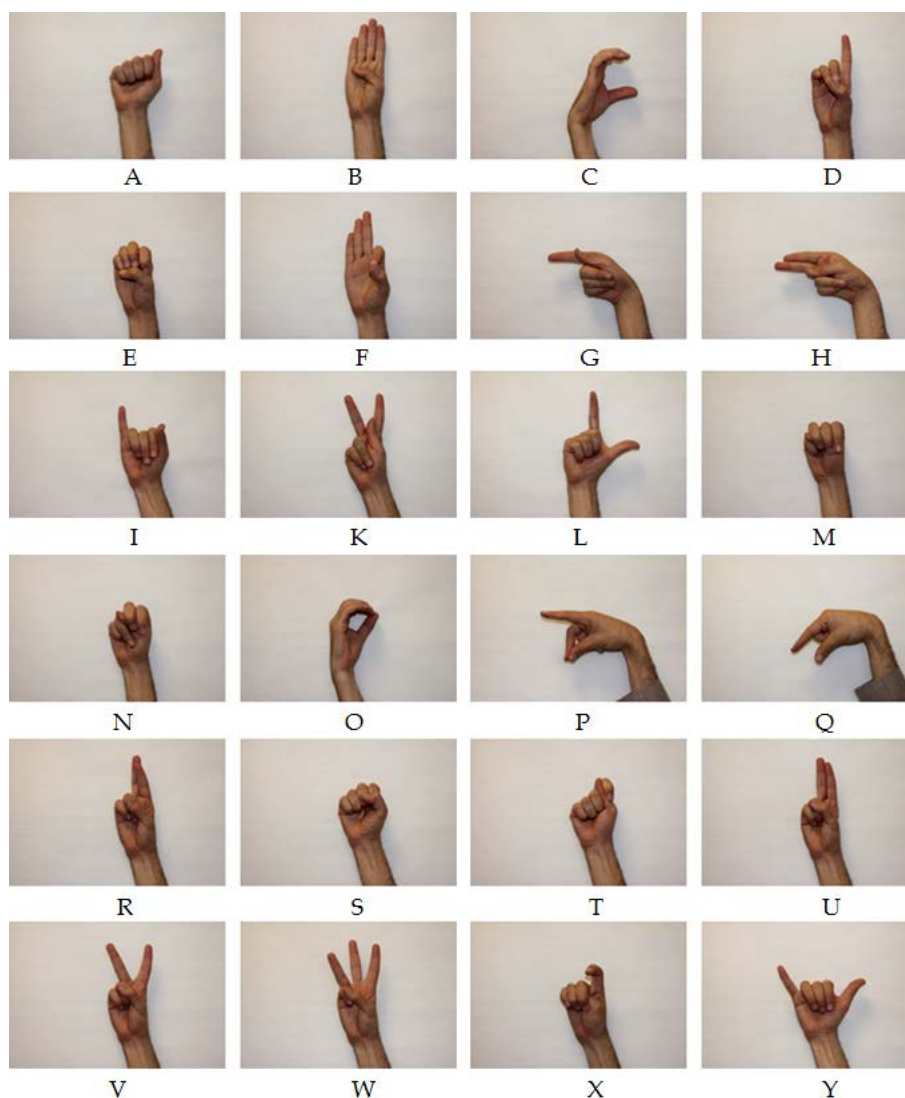
In order to express SL it usually involves the use of hands, arms, body movements and facial expressions; because of this, the development of a system to recognize all these expressions is a complex task; in fact most of the systems are very limited.

SLR process can be classified into two main classes based on how they acquire information; the first type uses Digital Image Processing, allowing users to interact in a more natural way with the system, but it is more difficult to acquire accurate data because most of these proposals work with 2D images, so it is difficult to follow the position or movement of the fingers or hand shape itself; on the other hand the second type SLRs use electronic devices physically connected to the user's body allowing to acquire accurate data of position, movement or velocity fingers and other points of interest, with the disadvantage of not letting free movement of users.

This is an extended work previously published in this journal [3]; the aim of this report is to recognize static signs by digital image processing (without movement sensors, special wires or any electronic device attached to signer) avoiding the use of special color markers or clothes.

## 2. Database

Mexican sign language consists in 26 signs, two of them are expressed by movement ("j" and "z"), for this reason the database created has 24 static signs (see [Figure 1](#)) were selected. A solid white background was used for segmentation purposes; images were captured by digital Canon EOS rebel T3 EF-S 18 - 55 camera using flash mode in order to decrease shadows. Five versions per sign were captured from single signer.



**Figure 1.** Static Mexican Sign Language (MSL) alphabets captured with a white background and avoiding the use of gloves or special color markers.

### 3. Jacobi-Fourier Moments

The technique named Jacobi-Fourier Moments [4] (JFM) is a powerful tool extensively used in image analysis. JFM is useful to extract relevant information from a function (in this case image of sign segmented in gray scale) and they are able to represent this function with few data with minimum redundancy due to its orthogonality property.

General expression of JFM is expressed as

$$A_{n,m} = \int_0^{2\pi} \int_0^1 f(r, \theta) P_{nm}(r, \theta) r dr d\theta, \quad (1)$$

where  $n$  denotes order and  $m$  repetition,  $f(r, \theta)$  is the image function in polar coordinates and  $P_{nm}(r, \theta)$  is the kernel function given by

$$P_{nm}(r, \theta) = J_n(\alpha, \beta, r) e^{jm\theta}, \quad (2)$$

where  $e^{jm\theta}$  is the Fourier term and  $J_n(\alpha, \beta, r)$  is the radial orthogonal Jacobi polynomial expressed as

$$J_n(\alpha, \beta, r) = \sqrt{\frac{w(\alpha, \beta, r)}{b_n(\alpha, \beta)}} G_n(\alpha, \beta, r). \quad (3)$$

$G_n(\alpha, \beta, r)$ ,  $w(\alpha, \beta, r)$  and  $b_n(\alpha, \beta)$  are Jacobi polynomials, weight function and normalization constant respectively and can be described using gamma function ( $\Gamma$ ) as [5]:

$$G_n(\alpha, \beta, r) = \frac{n! \Gamma(\beta)}{\Gamma(\alpha + n)} \sum_{s=0}^n (-1)^s \frac{\Gamma(\alpha + n + s)}{(n-s)! s! \Gamma(\beta + s)} r^s, \quad (4)$$

$$b_n(\alpha, \beta) = \frac{n! \Gamma^2(\beta) \Gamma(\alpha - \beta + n + 1)}{\Gamma(\beta + n) \Gamma(\alpha + n) (\alpha + 2n)} \quad (5)$$

and

$$w(\alpha, \beta, r) = (1-r)^{\alpha-\beta} r^{\beta-1}. \quad (6)$$

The restrictions for  $\alpha$  and  $\beta$  are  $\beta > 0$  and  $\alpha - \beta > -1$ .

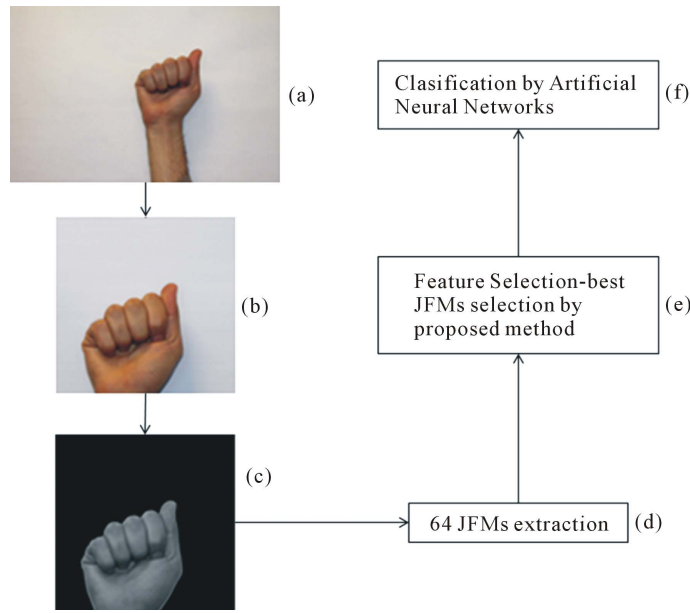
### 4. Proposed System

**Figure 2** shows the block diagram of proposed system. Original image can be seen in **Figure 2(a)**. In order to reduce computational costs a Region Of Interest (ROI) was selected by cutting the original image (see **Figure 2(b)**). **Figure 2(c)** illustrates the segmented alphabet “A” represented in gray scale which is used to calculate 64 JFM (Figure 2(d)). JFM are used as descriptors of signs, they use four parameters (order  $p$ , repetition  $q$ ,  $\alpha$  and  $\beta$ ). Experimentally we found best results in recognition rate for this database when  $\alpha = \beta = 1$ . 64 JFM were computed the combinations of  $p = 0, 1, 2, \dots, 7$  and  $q = 0, 1, 2, 3, \dots, 7$ , this features were also experimentally adjusted. Then the 64 JFM were sorted according to a metric that we propose that measures the performance of each JFM (process in **Figure 2(e)**). Finally 64 test are computed using a Multilayer Perceptron in WEKA [6], first test only uses the first JFM (best), second test uses first two sorted JFM, third test uses first three sorted JFM, and so on, the test number 64 uses all 64 JFM (process in **Figure 2(f)**).

The metric proposed to sort the JFM according to its performance in order to do a feature selection is described as follows. First a matrix  $[D]_{M \times N}$  of Descriptors (JFM) is defined to represent the JFM calculated in database, where  $M$  and  $n$  represent signs and versions respectively. A desirable JFM should be similar (numerically) when is calculated on different versions of same sign and should change when is calculated on different signs, this means that a JFM (with a particular  $\alpha, \beta, p$  and  $q$ ) computed in all database and represented by matrix  $[D]$  should be invariant along each row (low dispersion) and at the same time should be variant along each column (high dispersion).

In order to achieve a metric which considers the above mentioned some data are calculated. First a vector is computed to get the averages of versions per sign as

$$DN_i = \frac{1}{N} \sum_{j=1}^N D_{i,j} \quad i = 1, 2, \dots, M, \quad (7)$$



**Figure 2.** Block diagram of proposed system. (a) Original captured image of alphabet “A”; (b) cropped image; (c) segmented and RGB to gray scale converted; (d) 64 JFMs computed from “(c)”; (e) JFMs subset computed according to metric proposed and (f) database clasification using “(e)” and Multilayer Perceptron.

where  $DN_i$  stores the mean of versions per sign. This vector is used to calculate variance of versions as

$$SDN_i = \frac{1}{N} \sum_{j=1}^N (D_{i,j} - DN_i)^2 \quad i = 1, 2, \dots, M, \quad (8)$$

which is expected to be close to zero (minimum dispersion in versions), because descriptors should not change between the versions of same sign. Then versions average is calculated as

$$\overline{DN} = \frac{1}{M} \sum_{i=1}^M DN_i, \quad (9)$$

in order to compute variance of versions as

$$S\overline{DN} = \frac{1}{M} \sum_{i=1}^M (DN_i - \overline{DN})^2. \quad (10)$$

This two dispersion metrics ( $SDN_i$ -variance of versions and  $S\overline{DN}$ -variance of signs) are used to get a metric by a pondered ratio between them that estimates within a single value whether a JFM is good or not, this metric is expressed as

$$mo = \frac{1}{M} \sum_{i=1}^M \frac{S\overline{DN} + SDN_i}{S\overline{DN}}, \quad (11)$$

where  $mo$  is the objective metric that determines whether a JFM is good, this means that if  $S\overline{DN} \gg 0$  and  $SDN_i \cong 0$  then  $mo \cong 1$ , this value for  $mo$  is considered as desirable (minimum variance in versions and big variance in signs) for  $mo \gg 1$  means that is not a good descriptor due to  $V_i \gg 0$  or  $VS \cong 0$  or both (big variance in versions and/or minimum variance in signs).

64 JFMs were calculated and sorted by  $mo$  metric in ascending order then 64 tests were made in WEKA [6] using a Multilayer Perceptron (first introduced by Rosenblatt [7]). First test uses only a single descriptor to represent each image for all database, second test uses two descriptors, and so on. Last test uses all 64 descriptors to represent each image. Every test was made using cross validation.

**Table 1** shows the results of each classification test, first test which uses the JFM with  $p = 0$  and  $q = 0$

**Table 1.** 64 classification tests using a Multilayer Perceptron.

#	p	q	% classification	#	p	q	% classification
1	0	0	8.3333%	33	4	2	90.8333%
2	0	2	8.3333%	34	7	3	91.6667%
3	2	3	50.8333%	35	7	2	90.0%
4	0	3	60.0%	36	7	5	94.1667%
5	1	1	68.3333%	37	5	5	91.6667%
6	2	0	71.6667%	38	6	1	90.8333%
7	0	1	76.6667%	39	6	0	89.1667%
8	4	0	76.6667%	40	7	1	90.8333%
9	0	4	77.5%	41	5	6	91.6667%
10	1	0	80.0%	42	7	7	90.8333%
11	1	6	80.0%	43	3	7	91.6667%
12	0	5	82.5%	44	6	5	90.0%
13	1	2	87.5%	45	3	3	91.6667%
14	2	5	86.6667%	46	5	1	90.8333%
15	3	1	88.3333%	47	3	0	92.5%
16	1	3	87.5%	48	6	2	92.5%
17	5	0	86.6667%	49	1	4	90.8333%
18	2	6	88.3333%	50	7	4	90.0%
19	1	7	89.1667%	51	6	7	91.6667%
20	2	4	88.3333%	52	7	0	89.1667%
21	4	7	88.3333%	53	5	7	89.1667%
22	0	7	89.1667%	54	5	3	88.3333%
23	2	2	9.0%	55	1	5	90.0%
24	5	4	92.5%	56	6	6	90.0%
25	2	1	90.8333%	57	3	5	89.1667%
26	6	3	88.3333%	58	5	2	90.0%
27	4	6	<b>95.0%</b>	59	4	5	91.6667%
28	4	1	94.1667%	60	3	4	89.1667%
29	0	6	92.5%	61	4	4	92.5%
30	6	4	92.5%	62	3	6	90.8333%
31	3	2	92.5%	63	4	3	90.8333%
32	7	6	92.5%	64	2	7	89.1667%

achieves 8.3333% of recognition rate, second test uses two JFMs ( $p = 0$ ,  $q = 0$  and  $p = 0$ ,  $q = 2$ ) achieving 8.3333%, the best subset is the one which uses the first 27 JFMs to represent each image for all database achieving 95.0% of recognition rate.

## 5. Conclusion

JFMs can be used to extract descriptors of static signs. JFMs reduce computational cost for MSL recognition since an image can be represented by only 27 values. MSL recognition can be achieved without using gloves or special markers (using a special white background). A Multilayer Perceptron can be used to classify the signs using the JFMs and can achieve 95% of recognition rate in a cross validation scheme. The proposed metric can improve the global recognition rate; this can be seen in **Table 1** which shows that using all 64 JFMs 89.1667% of recognition rate was achieved and using the first 27 JFMs improves for this database the recognition rate in almost 6%.

## Acknowledgements

Authors thank to research department (Secretaría de Investigación) of Autonomous University of Mexico State (Universidad Autónoma del Estado de México) for the financial support to accomplish this work in the University Center UAEM of Teotihuacan Valley (Centro Universitario UAEM Valle de Teotihuacán) and the Polytechnic University of Tulancingo (Universidad Politécnica de Tulancingo) for all their support.

## References

- [1] Shohieb, S.M., Elminir, H.K. and Riad, A.M. (2015) SignsWorld Atlas; a Benchmark Arabic Sign Language Database. *Journal of King Saud University—Computer and Information Sciences*, **27**, 68-76.
- [2] López, V., Barra, R., Lutfi, S., Montero, J.M. and San, R. (2013) LSESpeak: Asproken Language Generator for Deaf People. *Expert Systems with Applications*, **40**, 1283-1295. <http://dx.doi.org/10.1016/j.eswa.2012.08.062>
- [3] Solís, F., Hernández, M., Pérez, A. and Toxqui, C. (2014) Static Digits Recognition Using Rotational Signatures and Hu Moments with a Multilayer Perceptron. *Engineering*, **6**, 692-698. <http://dx.doi.org/10.4236/eng.2014.611068>
- [4] Ping, Z., Ren, H., Zou, J., Sheng, Y. and Bo, W. (2007) Generic Orthogonal Moments: Jacobi-Fourier Moments for Invariant Image Description. *Pattern Recognition*, **40**, 1245-1254. <http://dx.doi.org/10.1016/j.patcog.2006.07.016>
- [5] Hoang, T. and Tabbone, S. (2013) Errata and Comments on “Generic Orthogonal Moments: Jacobi-Fourier Moments for Invariant Image Description”. *Pattern Recognition*, **46**, 3148-3155. <http://dx.doi.org/10.1016/j.patcog.2013.04.011>
- [6] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. (2009) The Weka Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, **11**, 10-18. <http://www.cs.waikato.ac.nz/ml/weka/> <http://dx.doi.org/10.1145/1656274.1656278>
- [7] Rosenblatt, F. (1962) Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. 1st Edition, Spartan Books, Michigan University, Ann Arbor.